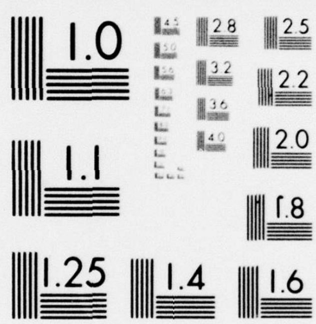
AD-A059 014 NORTH CAROLINA UNIV AT CHAPEL HILL DEPT OF BIOSTATISTICS F/G 12/1
INVARIANCE PRINCIPLES FOR JACKKNIFING U-STATISTICS FOR FINITE P--ETC(U)
1978 H MAJUMDAR, P K SEN AFOSR-74-2736

UNCLASSIFIED AFOSR-TR-78-1303 NL

1 OF 1
AD
A059014

END
DATE
FILMED
11-78
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 78-1303

(2) LEVEL II

AD A057014

DDC FILE COPY

INVARIANCE PRINCIPLES FOR JACKKNIFING U-STATISTICS FOR FINITE POPULATION SAMPLING AND SOME APPLICATIONS

Hiranmay Majumdar and Pranab Kumar Sen

University of North Carolina, Chapel Hill

DDC
RECEIVED
SEP 28 1978
B

ABSTRACT

For simple random sampling (without replacement) from a finite population, suitable stochastic processes are constructed from the entire sequence of jackknife estimators based on functions of U-statistics and these are approximated in distribution by some Brownian bridge processes. Strong convergence of the Tukey estimator of the variance of jackknife U-statistics has also been established. Some applications of these results in sequential analysis relating to finite population sampling are also considered.

1. INTRODUCTION

Let Ω_N be a finite population of size N, represented by the vector $\underset{\sim}{A}_N = (a_{N1},\ldots,a_{NN})$ of real numbers. Let $\underset{\sim}{X}_N = (X_{N1},\ldots,X_{NN})$ be a random vector which takes on each permutation of the elements of $\underset{\sim}{A}_N$ with equal probability $(N!)^{-1}$. Then a random sample of size $n(\leq N)$ drawn without replacement from Ω_N may be represented by $\underset{\sim}{X}_{Nn} = (X_{n1},\ldots,X_{Nn})$, so that $\underset{\sim}{X}_{Nn}$ takes

78 09 13 088

Approved for public release; distribution unlimited.

on all possible n-tuples $(a_{Ni_1},\dots,a_{Ni_n})$, $i \le i_1 \ne \dots \ne i_n \le N$, with the common probability $N^{-[n]}\{(= N\dots(N-n+1)\}^{-1})$, for $n = 1,\dots,N$.

For a *symmetric kernel* $f(\underset{\sim}{X}_{Nm})$ of *degree* $m(\ge 1)$, the U-statistic, defined by

$$U_{Nn} = U(\underset{\sim}{X}_{Nn}) = n^{-[m]}\sum_{P_{n,m}} f(X_{Ni_1},\dots,X_{Ni_m}),\ n \ge m, \tag{1.1}$$

(where $P_{n,m} = \{(i_1,\dots,i_m): 1 \le i_1 \ne \dots \ne i_m \le n\}$), is an unbiased estimator of

$$\mu_N = U(\underset{\sim}{X}_{NN}) = U(\underset{\sim}{A}_N) = N^{-[m]}\sum_{P_{N,m}} f(a_{Ni_1},\dots,a_{Ni_m}). \tag{1.2}$$

Various properties of U_{Nn} have been studied by Nandi and Sen (1963), Sen (1970, 1972) and others.

Let us consider a real-valued function

$$\theta_N = g(\mu_N), \tag{1.3}$$

where g is a smooth function. Though U_{Nn} is an unbiased estimator of μ_N, $\hat\theta_{Nn} = g(U_{Nn})$ is not generally unbiased for θ_N. For this reason, we consider the following *jackknife estimator*. Let

$$U_{Nn-1}^{(i)} = (n-1)^{-[m]}\sum_{P^i_{n-1,m}} f(X_{Ni_1},\dots,X_{Ni_m}), \tag{1.4}$$

where $P^i_{n-1,m} = \{(i_1,\dots,i_m): 1 \le i_1 \ne \dots \ne i_m \le n$ with $i_j \ne i$, $1 \le j \le m\}$, for $i = 1,\dots,n$. Also, let

$$\hat\theta^{(i)}_{Nn-1} = g(U^{(i)}_{Nn-1}),\ 1 \le i \le n; \tag{1.5}$$

$$\hat\theta_{Nn,i} = n\hat\theta_{Nn} - (n-1)\hat\theta^{(i)}_{Nn-1},\ 1 \le i \le n; \tag{1.6}$$

$$\theta^*_{Nn} = n^{-1}\sum_{i=1}^{n}\hat\theta_{Nn,i}. \tag{1.7}$$

Then θ^*_{Nn} is the jackknife estimator of θ_N.

For random sampling from an infinite population, jackknifing of U-statistics has been studied by Arvesen (1969). Recently, Sen (1977) has carried the investigation further by establishing invariance principles for jackknife statistics and incorporating

78 09 13 088

them to some problems in sequential analysis. The object of the present investigation is to extend the results of Sen (1977) to sampling from finite population and to emply them in some problems of survey sampling.

The basic assumptions and preliminary notions are outlined in Section 2. Section 3 deals with the main theorems and their derivations. In Section 4, some applications and clarification of certain results, which have so far been tacitly assumed by the workers in this field, are discussed.

2. PRELIMINARY NOTIONS

As in Sen (1972), we define for h: $0 \le h \le m$,

$$f_h(X_{i_1},\dots,X_{i_h}) = (N-h)^{-[m]}\sum\nolimits_{(h)} f(X_{i_1},\dots,X_{i_m}) \,, \tag{2.1}$$

where $\sum_{(h)}$ extends over all $1 \le i_{h+1} \ne \dots \ne i_m \le N$ with $i_{h+j} \ne i_s$ for $j = 1,\dots,m-h$ and $s = 1,\dots,h$. Then $f_0 = \mu_N$ and $f_m = f$. Also, let

$$\begin{aligned}\bar{\zeta}_{h,N} &= \mathrm{Var}\{f_h(\underset{\sim}{X}_{Nh})\} \\ &= N^{-[h]}\sum\nolimits_{P_{N,h}} f_h^2(a_{Ni_1},\dots,a_{Ni_h}) - \mu_N^2 \,,\end{aligned} \tag{2.2}$$

for $0 \le h \le m$, where $\bar{\zeta}_{0,N} = 0$ and it follows from Nandi and Sen (1963) that

$$0 \le \bar{\zeta}_{h,N} \le (h/g)\bar{\zeta}_{g,N} \,,\ \forall\ 1 \le h \le g \le m \,. \tag{2.3}$$

For the study of asymptotic properties, we conceive of a sequence $\{\Omega_N\}$ of populations and allow $N \to \infty$. We assume that

(A) $\inf\limits_N \bar{\zeta}_{1,N} > 0$ and $\sup\limits_N \bar{\zeta}_{m,N} < \infty$ (2.4)

(B) $\sup\limits_N E|f(\underset{\sim}{X}_{Nm})|^4 < \infty$ (2.5)

and (C) g, in (1.3), has a bounded second derivative in some neighborhood of μ_N. Note that the second condition in (2.4) follows from (2.5).

Note that by (3.22) of Nandi and Sen (1963), $\forall\ N \ge n \ge m$,

$$n^{-1}m^2\left\{\binom{N-n}{m}\Big/\binom{N-m}{m}\right\}\bar{\zeta}_{1,N} \le V\{U_{Nn}\} \le n^{-1}m\left\{\frac{N-n}{N-2m+1}\right\}\bar{\zeta}_{m,N} \,, \tag{2.6}$$

'N for
White Section ☑
Buff Section ☐
'NCED ☐
'ION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

so that by (2.4), for $m \le n \le N(1-\varepsilon)$, $\varepsilon > 0$,

$$0 < \inf_N nV(U_{Nn}) \le \sup_N nV(U_{Nn}) < \infty , \qquad (2.7)$$

while $nV(U_{Nn}) \to 0$ as $n \to N$. Further, by (2.4) and (2.5),

$$\max_{1\le i\le N} \{|f_1(a_{Ni}) - \mu_N|/[\zeta_{1,N}^{-\frac{1}{2}}]\} = O(N^{1/4}) , \quad \text{as} \quad N \to \infty . \qquad (2.8)$$

Let us also define, as in Sen (1972),

$$U_{Nn}^{(h)} = n^{-[h]}\sum_{P_{n,h}} f_h(X_{i_1},\dots,X_{i_h}) , \quad 0 \le h \le m , \qquad (2.9)$$

where $U_{Nn}^{(0)} = \mu_N$ and $U_{Nn}^{(m)} = U_{Nn}$, $\forall\, n \le m$, and let

$$W_{Nn}^{(h)} = \sum_{k=0}^{h} \binom{h}{k}(-1)^k U_{Nn}^{(h-k)} , \quad h = 0,1,\dots,m . \qquad (2.10)$$

Then, as in Sen (1972),

$$U_{Nn} = \sum_{h=0}^{m} \binom{m}{h} W_{Nn}^{(h)} . \qquad (2.11)$$

Given that the collection $(a_{Ni_1},\dots,a_{Ni_n})$ corresponds to a sample $\underset{\sim}{X}_{Nn}$ (without specifying the order in which the elements occur), $\underset{\sim}{X}_{Nn}$ can assume any one of the $n!$ possible permutations of $(a_{Ni_1},\dots,a_{Ni_n})$ with the same conditional probability $(n!)^{-1}$. The conditional expectation with respect to this conditional distribution is denoted by $E(\cdot|C_{Nn})$. Then,

$$E(\hat{\theta}_{Nn-1}|C_{Nn}) = n^{-1}\sum_{i=1}^{n} \hat{\theta}_{Nn-1}^{(i)} , \quad \forall\, n > m , \qquad (2.12)$$

and, as a result, by (1.5)-(1.7) and (2.12),

$$\theta_{Nn}^{*} = \hat{\theta}_{Nn} + (n-1)E\{(\hat{\theta}_{Nn} - \hat{\theta}_{Nn-1})|C_{Nn}\} , \quad \forall\, n > m . \qquad (2.13)$$

The Tukey estimator of the variance of $n^{\frac{1}{2}}(\theta_{Nn}^{*}-\mu_N)[(N-n)/N]^{-\frac{1}{2}}$ is

$$V_{Nn} = (n-1)^{-1}\sum_{i=1}^{n}(\hat{\theta}_{Nn,i} - \theta_{Nn}^{*})^2 , \qquad (2.14)$$

where by similar arguments if follows that

$$V_{Nn} = n(n-1)\ \mathrm{Var}\{(\hat{\theta}_{Nn} - \hat{\theta}_{Nn-1})|C_{Nn}\} . \qquad (2.15)$$

Both (2.13) and (2.15) are in agreement with the parallel results for infinite populations, treated in Sen (1977). We conclude this section with the following.

Definition: *Let* $\{T_{Nn}, n \le N, N \le 1\}$ *be a double sequence of statistics and* $\{\alpha_N, N \ge 1\}$ *be a sequence of real numbers. Then,* $T_{Nn} - \alpha_N$ *strongly converges (s.c.) to* 0, *if for every* $\varepsilon > 0$ *and every sequence* $\{N^*(\le N)\}$, *such that* $N^* \to \infty$ *as* $N \to \infty$ *(but* N^*/N *may or may not go to* 0*)*,

$$\overline{\lim}\, P\left\{ \max_{N^* \le n \le N} |T_{Nn} - \alpha_N| > \varepsilon \right\} = 0 \ . \tag{2.16}$$

We shall find this definition useful in the subsequent sections.

3. INVARIANCE PRINCIPLES FOR $\{\theta^*_{Nn}\}$

Let $\{N^*\}$ be a sequence of positive integers such that as $N \to \infty$, $N^* \to \infty$ but $N^*/N \to 0$ (viz., $N^* = N^\lambda$, $0 < \lambda < 1$ of $\log N$ etc.). We then consider the stochastic process $Z_N = \{Z_N(t), t \in I = [0,1]\}$ by letting $t_{N,k} = k/N$, $k = 0,1,\dots,N$ and

$$Z_N(t) = \begin{cases} 0\ , & t < N^*/N\ , \\ Z_{Nk} = k(NV_{Nk})^{-\frac{1}{2}}(\theta^*_{Nk} - \theta_N)\ , & t = t_{N,k}\ , \\ (k+1-t)Z_{Nk} + (t-1)Z_{Nk+1}\ , & t_{N,k} \le t \le t_{N,k+1}\ , \end{cases} \tag{3.1}$$

for $k \le N$. Then, Z_N has a continuous sample path and it belongs to the space $C[0,1]$ of continuous real-valued functions on I. Let $Z^0 = \{Z^0(t)\ t \in I\}$ be a standard Brownian bridge on I. That is $\underset{\sim}{Z}^0$ is a Gaussian function with $EZ^0(t) = 0$ and $EZ^0(s)Z^0(t) = s \wedge t - st = \min(s,t) - st$, $\forall\ s,t \in I$. We say [viz., Billingsley (1968)] that Z_N converges in law (or distribution) to Z^0 if for every continuous functional $h(\cdot)$ assuming values in R^k, the $k(\ge 1)$-dimensional Euclidean space, as $N \to \infty$, $h(Z_N)$ has asymptotically the same distribution as of $h(Z)$. For example, the above weak convergence of Z_N to Z^0 insures that $\sup\{Z_N(t):\ 0 \le t \le 1\}$, $\sup\{|Z_N(t)|:\ 0 \le t \le 1\}$ and $\int_0^1 Z_N^2(t)dt$ have the same limiting distributions as of $\sup\{Z^0(t):\ 0 \le t \le 1\}$, $\sup\{|Z^0(t)|:\ 0 \le t \le 1\}$ and $\int_0^1 [Z^0(t)]^2 dt$, respectively. This mode of convergence is stronger than the asymptotic normality of $n^{\frac{1}{2}}(\theta^*_{Nn} - \theta)$ and it also insures that for finitely many

$t_1, \ldots, t_m$ (all belong to I), $[Z_N(t_1), \ldots, Z_N(t_m)]$ has asymptotically a multinormal distribution. Then, we have the following.

Theorem 3.1. *Under Assumptions* (A), (B) *and* (C) *of Section* 2, Z_N *converges in law to* Z^0.

Before we present the proof of the theorem, we consider several results. First, the following theorem (whose proof is postponed to Section 5) is of basic importance in this context.

Theorem 3.2. *Under* (2.5), $|n(n-1)E\{(U_{Nn-1}-U_{Nn})^2|C_{Nn}\} - m^2\bar{\zeta}_{1,N}| \to 0$ s.c., *as* $N \to \infty$.

Note that $n(n-1)E\{(U_{Nn-1}-U_{Nn})^2|C_{Nn}\} = (n-1)\sum_{i=1}^{n}(U_{Nn-1}^{(i)}-U_{Nn})^2$ $\left(\geq (n-1)\left\{\max_{1\leq i\leq n}(U_{Nn-1}^{(i)}-U_{Nn})^2\right\}\right)$, so that by (2.5) and Theorem 3.2,

$$\max_{N^*\leq n\leq N}\ \max_{1\leq i\leq n}(n-1[U_{Nn-1}^{(i)}-U_{Nn}]^2 = 0(1), \text{ in probability} . \quad (3.2)$$

Further, by Theorem 1 of Sen (1970), $U_{Nn} - \mu_N \to 0$ s.c., and hence, by (3.2),

$$\max_{1\leq i\leq n}|U_{Nn-1}^{(i)}-\mu_N| \to 0 \text{ s.c., as } N \to \infty . \quad (3.3)$$

Let us now define

$$\gamma_N^2 = [g'(\mu_N)]^2 m^2\bar{\zeta}_{1,N}, \quad N \geq 1 . \quad (3.4)$$

Then, by virtue of Theorem 3.2, and (3.2) and (3.3), we may virtually repeat the steps in the proof of Theorem 3.1 of Sen (1977) and obtain that under Assumptions (B) and (C),

$$V_{Nn} - \gamma_N^2 \to 0 \text{ s.c., as } N \to \infty$$

$$\theta^*_{Nn} - \hat{\theta}_{Nn} = 0(n^{-1}) \text{ in the strong sense in (2.16)} . \quad (3.6)$$

Let us now return to the proof of Theorem 3.1. Suppose that in (3.1), we replace $V_{Nk_N(t)}$ by γ_N^2 and denote the resulting process by $Z_N^0 = \{Z_N^0(t), t \in I\}$. Then, by definition,

$$\begin{aligned}\rho(Z_N, Z_N^0) &= \sup_{t\in I}|Z_N(t) - Z_N^0(t)| \\ &= \sup_{t\in I^*}\{|\gamma_N/V_{Nk_n(t)}^{1/2} - 1||Z_N^0(t)|\} \quad \left(I^* = \left[\frac{N^*}{N}, 1\right]\right) . \quad (3.7)\end{aligned}$$

Hence, if Z_N^0 converges weakly to Z^0 (which implies that $\sup\{|Z_N^0(t)|: t \in I\} = O_p(1)$), then, by (3.5) and (3.7), $\rho(Z_N, Z_N^0) \xrightarrow{p} 0$ as $N \to \infty$. Hence, it suffices to prove the following

<u>Theorem 3.3.</u> *Under the hypothesis of Theorem 3.1, Z_N^0 converges in law to Z^0.*

<u>Proof.</u> We note that

$$\begin{aligned}\hat{\theta}_{Nn} - \theta_N &= g(U_{Nn}) - g(\mu_N) \\ &= g'(\mu_N)[U_{Nn} - \mu_N] \\ &\quad + \tfrac{1}{2}[U_{Nn} - \mu_N]^2 g''(hU_{Nn} + (1-h)\mu_N),\ 0<h<1, \end{aligned} \tag{3.8}$$

where by Assumption (C), $g''(.)$ is bounded in some neighborhood of μ_N. Also, by Theorem 1 of Sen (1970), $n^{\frac{1}{2}}(U_{Nn} - \mu_N)^2 \to 0$ s.c., and hence,

$$nN^{-\frac{1}{2}}|(\hat{\theta}_{Nn} - \theta_N) - g'(\mu_N)[U_{Nn} - \mu_N]| \to 0 \text{ s.c.} \tag{3.9}$$

Suppose now that in (3.1), we replace $\theta^*_{Nk} - \theta$ and V_{Nk} by $g'(\mu_N)(U_{Nk} - \mu_N)$ and γ_N^2 , respectively, and denote the resulting process by $\tilde{Z}_N = \{\tilde{Z}_N(t), t \in I\}$. Then, by Theorem 2.1 of Sen (1972), $\tilde{Z}_N$ converges in law to Z^0, while by (3.6) and (3.9), along with the weak convergence of $\tilde{Z}_N$,

$$\rho(Z_N^0, \tilde{Z}_N) \xrightarrow{p} 0, \tag{3.10}$$

and hence, Z_N^0 converges weakly to Z^0. Q.E.D.

So far, we have assumed that the a_{Ni} are all real numbers. There is no harm in letting them be real p-vectors, for some $p \geq 1$. The same *permutation argument* holds in this case, and hence, the proofs remain unaltered. Also, in practical problems, $\underset{\sim}{U}_{Nn} = (U_{Nn(1)}, \ldots, U_{Nn(q)})$ (for some $q \geq 1$) may be a q-vector, where the $U_{Nn(j)}$ are defined by (1.1) for kernels $f_{(j)}$ of degree $m_j (\geq 1)$ and in the same framework, we have $\underset{\sim}{\mu}_N = \underset{\sim}{U}_{NN}$, $\theta_N = g(\underset{\sim}{\mu}_N)$ and $\hat{\theta}_{Nn} = g(\underset{\sim}{U}_{Nn})$, where we assume that g has bounded second order partial derivatives in some neighborhood of $\underset{\sim}{\mu}_N$. Then, *replacing in* (1.5), $U_{Nn-1}^{(i)}$ by $\underset{\sim}{U}_{Nn-1}^{(i)}$ [defined as in (1.4) with

$\underset{\sim}{f} = (f_{(1)}, \ldots, f_{(q)})]$, we define the jackknife estimator θ^{*}_{Nn} as in (1.6)-(1.7). If we define the $f_{(j)h}$ as in (2.1) for $f = f_{(j)}$ and replace, in (2.2), f_h^2 by $f_{(j)h}(\underset{\sim}{X}_{Nh}) f_{(\ell)h}(\underset{\sim}{X}_{Nh})$, the resulting quantity is denoted by $\bar{\zeta}_{h(j\ell),N}$ for $j,\ell = 1,\ldots,p$ and $h \geq 0$. Then, assuming that (2.4)-(2.5) hold for each $j(=1,\ldots,q)$, it follows by arguments very similar to the ones in the proof of Theorem 3.2 (see Appendix) that for the jackknife variance V_{Nn},

$$V_{Nn} - \gamma_N^2 \to 0 \quad \text{s.c.} , \tag{3.11}$$

where

$$\gamma_N^2 = \sum_{j=1}^{q} \sum_{\ell=1}^{q} g_j'(\underset{\sim}{\mu}_N) g_\ell'(\underset{\sim}{\mu}_N) m_j m_\ell \bar{\zeta}_{1(j\ell),N} . \tag{3.12}$$

With this modification, Theorems 3.1 and 3.2 hold under no extra regularity condition. Actually, (3.2) and (3.3) hold coordinatewise for each $j(=1,\ldots,q)$, in (3.8) we have a multivariate second order Taylor series expansion, so that in (3.9) (and in the definition of $\tilde{Z}_N$), $g'(\mu_N)(U_{Nn}-\mu_N)$ has to be replaced by $\sum_{j=1}^{q} g_j'(\underset{\sim}{\mu}_N)(U_{Nn(j)} - \mu_{N(j)})$ which, being a linear combination of $\underset{\sim}{U}_{Nn} - \underset{\sim}{\mu}_N$, is itself a U-statistic, and hence, Theorem 2.1 of Sen (1972) holds directly.

It is also possible to consider a vector $\underset{\sim}{\theta}_N = \underset{\sim}{g}(\underset{\sim}{U}_{NN})$ and $\hat{\underset{\sim}{\theta}}_{Nn} = \underset{\sim}{g}(\underset{\sim}{U}_{Nn})$, where $\underset{\sim}{g}(\cdot) = (g_{(1)}(\cdot), \ldots, g_{(r)}(\cdot))$ for some $r \geq 1$ and $\underset{\sim}{U}_{Nn} = (U_{Nn(1)}, \ldots, U_{Nn(q)})$, for some $q \geq 1$. In such a case, we assume that for each $g_{(s)}$ and $U_{Nn(j)}$, the Assumptions (A), (B) and (C) of Section 2 are met. Defining the $\underset{\sim}{U}^{(i)}_{Nn-1}$ by (1.4) for $\underset{\sim}{f} = (f_{(1)}, \ldots, f_{(q)})$ and $\hat{\underset{\sim}{\theta}}^{(i)}_{Nn-1} = \underset{\sim}{g}(\underset{\sim}{U}^{(i)}_{Nn-1})$, $1 \leq i \leq n$, the jackknife estimator $\underset{\sim}{\theta}^{*}_{Nn}$ is again defined by (1.6)-(1.7). The Tukey estimator of the dispersion matrix of $n^{\frac{1}{2}}(\underset{\sim}{\theta}^{*}_{Nn} - \underset{\sim}{\theta}_N)$ is given by

$$\underset{\sim}{V}_{Nn} = (n-1)^{-1} \sum_{i=1}^{n} (\hat{\underset{\sim}{\theta}}_{Nn,i} - \underset{\sim}{\theta}^{*}_{Nn})'(\hat{\underset{\sim}{\theta}}_{Nn,i} - \underset{\sim}{\theta}^{*}_{Nn}) . \tag{3.13}$$

Let us consider the matrix $\underset{\sim}{\Gamma}_N = ((\gamma_{N,ss'}))$ where, for every $s,s'(=1,\ldots,r)$,

$$\gamma_{N,ss'} = \sum_{j=1}^{q} \sum_{\ell=1}^{q} m_j m_\ell g'_{(s)j}(\underset{\sim}{\mu}_N) g'_{(s')\ell}(\mu_N) \bar{\zeta}_{1(j\ell),N} \tag{3.14}$$

and $g'_{(s)j}$ is the partial derivative of $g_{(s)}$ with respect to the j-th argument, for $j = 1,\ldots,q$ and $s = 1,\ldots,r$. Then, by a direct coordinatewise extension of (3.11)-(3.12) we have

$$\underset{\sim}{V}_{Nn} - \underset{\sim}{\Gamma}_N \to 0 \quad \text{s.c.} \tag{3.15}$$

Consider then a vector-valued stocastic process $\underset{\sim}{Z}_N = \{\underset{\sim}{Z}_N(t),\ t \in I\}$ where $\underset{\sim}{Z}_N(t)$ is defined [as in (3.1)] by linear interpolation of $\underset{\sim}{Z}_{Nk}$, $0 \le k \le N$ and

$$\underset{\sim}{Z}_{Nk} = \begin{cases} kN^{-\frac{1}{2}}\underset{\sim}{V}_{Nk}^{-\frac{1}{2}}(\underset{\sim}{\theta}^{\star}_{Nk} - \underset{\sim}{\theta}_N) & \text{if } \underset{\sim}{V}_{Nk} \text{ is positive definite} \\ & \text{and } N^{\star} \le k \le N\,, \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3.16}$$

Finally, let $\underset{\sim}{Z}^0 = \{\underset{\sim}{Z}^0(t),\ t \in I\}$ be a vector-Gaussian function on I with $E\underset{\sim}{Z}^0(t) = \underset{\sim}{0}$ and $E(\underset{\sim}{Z}^0(t))'(\underset{\sim}{Z}^0(s)) = \{(s \wedge t) - st\}\underset{\sim}{I}_r$, $\forall\ s, t \in I$, where $\underset{\sim}{I}_r$ is the unit matrix of order r. Thus, the components of $\underset{\sim}{Z}^0$ are all independent Brownian bridges on I. Then, we have the following.

Theorem 3.4. *Under the conditions mentioned above,* $\underset{\sim}{Z}_N$ *converges in law to* $\underset{\sim}{Z}^0$, *whenever* $\underset{\sim}{\Gamma}$ *is p.d.*

Outline of the proof. Let us define $\underset{\sim}{Z}^0_N = \{\underset{\sim}{Z}^0_N(t),\ t \in I\}$, by replacing in (3.16) $\underset{\sim}{V}_{Nk_N(t)}$ by $\underset{\sim}{\Gamma}_N$. Then, by arguments similar to those in (3.7),

$$\rho(\underset{\sim}{Z}_N, \underset{\sim}{Z}^0_N) \le p\left[\sup_{t \in I^{\star}}\{||\underset{\sim}{\Gamma}_N\underset{\sim}{V}^{-1}_{Nk_N(t)} - \underset{\sim}{I}_r||\ ||\underset{\sim}{Z}^0_N(t)||\}\right]\,, \tag{3.17}$$

where $||\underset{\sim}{a}|| = (\underset{\sim}{a}\underset{\sim}{a}')^{\frac{1}{2}}$ and $||\underset{\sim}{A}||^2$ = trace of $\underset{\sim}{A}^2$. Hence, by virtue of (3.15), it suffices to show that $\underset{\sim}{Z}^0_N \overset{\mathcal{D}}{\to} \underset{\sim}{Z}^0$. Towards this, we consider a direct (coordinatewise) extension of (3.8)-(3.9) [with modifications as in after (3.12)], and the rest of the proof follows by using a direct (vector-) extension of Theorem 2.1 of Sen (1972).

4. APPLICATIONS

We conceive of a random sample $(\underset{\sim}{X}_{Nn})$ of size n drawn without replacement from Ω_N where the $\underset{\sim}{a}_{Ni}$ (and hence, X_{Ni}) are

all p-vectors, for some $p \geq 1$. We consider the following applications.

4.1. Estimation and Testing of Multiple Regression Coefficients

Let us denote the population dispersion matrix by

$$\underset{\sim}{\Lambda}_N = N^{-1} \sum_{i=1}^{N} (\underset{\sim}{a}_{Ni} - \bar{\underset{\sim}{a}}_N)'(\underset{\sim}{a}_{Ni} - \bar{\underset{\sim}{a}}_N) \; ; \; \bar{\underset{\sim}{a}}_N = \frac{1}{N} \sum_{i=1}^{N} \underset{\sim}{a}_{Ni} \; . \tag{4.1}$$

We write $\underset{\sim}{\Lambda}_N = ((\lambda_{Nrs}))_{r,s=1,\ldots,p}$ and denote the minor of λ_{Nrs} by $\underset{\sim}{\Lambda}_{Nrs}$, $r,s=1,\ldots,p$. Also, we denote X_{Ni} by $(X_{Ni}^{(i)},\ldots,X_{Ni}^{(p)})$, $i = 1,\ldots,N$. Then the population regression coefficients of $X_{Ni}^{(p)}$ on $(X_{Ni}^{(1)},\ldots,X_{Ni}^{(p-1)})$ $(= \underset{\sim}{X}_{Ni}^{*}$, say) are given by

$$\underset{\sim}{\beta}_N = \underset{\sim}{\Lambda}_{Npp}^{-1}\underset{\sim}{\lambda}_{Np}^{*} \quad \text{where} \quad \underset{\sim}{\lambda}_{Np}^{*} = (\lambda_{Np1},\ldots,\lambda_{Npp-1})', \tag{4.2}$$

The usual sample estimator of $\underset{\sim}{\beta}_N$ (based on $\underset{\sim}{X}_{Nn}$) is

$$\underset{\sim}{b}_{Nn} = \underset{\sim}{L}_{Nnpp}^{-1}\underset{\sim}{\ell}_{Nnp}^{*} \quad \text{where} \quad \underset{\sim}{\ell}_{Nnp}^{*} = (\ell_{Nnp1},\ldots,\ell_{Nnpp-1})', \tag{4.3}$$

$\underset{\sim}{L}_{Nn} = ((\ell_{Nnrs}))$, $\underset{\sim}{L}_{Nnrs}$ is the minor of ℓ_{Nnrs} and

$$\underset{\sim}{L}_{Nn} = n^{-[2]}\sum_{1\leq i\neq j\leq n} \underset{\sim}{f}(X_{Ni},X_{Nj}) \tag{4.3}$$

and

$$\underset{\sim}{f}(X_{Ni},X_{Nj}) = \tfrac{1}{2}\,(X_{Ni}-X_{Nj})'(X_{Ni}-X_{Nj}) \tag{4.4}$$

is a matrix of order $p \times p$. Though $\underset{\sim}{L}_{Nn}$ is unbiased for $\underset{\sim}{\Lambda}_N$, $\underset{\sim}{b}_{Nn}$ is not necessarily so. But $\underset{\sim}{L}_{Nn}$ is a matrix of U-statistics, and we are natrually tempted to use jackknifing to reduce the bias of $\underset{\sim}{b}_{Nn}$; the jackknife estimator as defined in (1.7) [and before (3.13)] is denoted by $\underset{\sim}{b}_{Nn}^{*}$. Using the results of MacRae (1974), we have

$$\frac{\partial \underset{\sim}{b}_{Nn}}{\partial \underset{\sim}{L}_{Nnpp}} = -\{\underset{\sim}{L}_{Nnpp}^{-1} \otimes \underset{\sim}{I}_{p-1}\}\underset{\sim}{E}(p-1,p-1)\{\underset{\sim}{L}_{Nnpp}^{-1} \otimes I_{p-1}\}\underset{\sim}{\ell}_{Nnp}^{*} \tag{4.5}$$

where $\otimes$ denotes the direct product,

$$\underset{\sim}{E}(p-1,p-1) = ((\underset{\sim}{E}_{ij}))_{i,j=1,\ldots,p-1} \;\; (\text{of order } (p-1)^2 \times (p-1)^2), \tag{4.6}$$

and $\underset{\sim}{E}_{ij}$ is a $(p-1)\times(p-1)$ matrix with the element 1 in its (i,j) and (j,i)-th positions and 0 elsewhere, for $i,j = 1,\ldots,p-1$. Also,

$$\frac{\partial \underset{\sim}{b}_{Nn}}{\partial \underset{\sim}{\ell}^{\star}_{Nnp}} = \underset{\sim}{L}^{-1}_{Nnpp}\underset{\sim}{1}_{p-1} . \tag{4.7}$$

Thus, if we assume that (i) the characteristic roots of $\underset{\sim}{\Lambda}_N$ are all bounded away from 0 and ∞, then condition (C) of Section 2 holds, while (2.5) holds, if in addition, we assume that

$$\overline{\lim}\; N^{-1} \sum_{i=1}^{N} [a_{Ni}^{(j)} - \bar{a}_N^{(j)}]^8 < \infty, \quad \forall\; j = 1,\ldots,p . \tag{4.8}$$

Let $\underset{\sim}{V}_{Nn}$ be the jackknife dispersion matrix of $n^{\frac{1}{2}}(\underset{\sim}{b}^{\star}_{Nn} - \underset{\sim}{\beta}_N)$, defined by (3.13). Then from Theorem 3.4, we claim that whenever $n(< N)$ is large, $T_{Nn} = nN(N-n)^{-1}(\underset{\sim}{b}^{\star}_{Nn} - \underset{\sim}{\beta}_N)'\underset{\sim}{V}^{-1}_{Nn}(\underset{\sim}{b}^{\star}_{Nn} - \underset{\sim}{\beta}_N)$ has asymptotically the chi-square distribution with $p-1$ degrees of freedom (DF) and this can be incorporated in testing a null hypothesis H_0: $\underset{\sim}{\beta}_N = \underset{\sim}{\beta}^0$ (specified) against $H^{\star}$: $\underset{\sim}{\beta}_N \neq \underset{\sim}{\beta}^0$ or to provide a (simultaneous) confidence region for $\underset{\sim}{\beta}_N$. Let $\chi^2_{t,\alpha}$ be the upper $100\alpha\%$ point of the chi-square distribution with t DF and let T^0_{Nn} be defined by replacing $\underset{\sim}{\beta}_N$ by $\underset{\sim}{\beta}^0$ in the definition of T_{Nn}. Then, we accept or reject H_0 according as T^0_{NN} is $\leq$ or $> \chi^2_{p-1,\alpha}$ where $\alpha(0 < \alpha < 1)$ is the desired level of significance of the test. Further, if we consider the (ellipsoidal) region

$$I_{Nn} = \{\underset{\sim}{\beta}_N : T_{Nn} \leq \chi^2_{p-1,\alpha}\} \tag{4.9}$$

the for large n, I_{Nn} provides a confidence region for $\underset{\sim}{\beta}_N$ with confidence coefficient $1-\alpha$.

In actual practice, neither $\underset{\sim}{V}_{Nn}$ nor its population counterpart $\underset{\sim}{\Gamma}_N$, defined by (3.14), is usually known in advance, and hence, an improper choice of n may result either in excessive costing (due to over-sampling) or in a larger diameter of I_{Nn} [or small power of the test based on T^0_{Nn}] (due to under-sampling). For this reason, a sequential procedures may be adapted which through regular updating of information through sampling results in (nearly) optimal solutions. These sequential procedures, in turn, rests on the invariance principles considered in Section 3. Let $ch_1(\underset{\sim}{B})$ be the largest characteristic root of $\underset{\sim}{B}$ and let

$$n_d^* = \min\{k: k \geq n^0 \text{ and } ch_1(\underset{\sim}{V}_{Nk}) \leq d^2 kN[(N-k)\chi^2_{p-1,\alpha}]^{-1}\} , \quad (4.10)$$

where $d(>0)$ is a preassigned (small) positive number and $n^0(>p)$ is an initial sample size with which sampling commences. Thus n_d^* is a positive integer valued random variable and $n^0 \leq n_d^* \leq N$. Then, starting with the sample size n^0, units are drawn (one by one) without replacement so long as $k \leq n_d^*$ i.e., $ch_1(\underset{\sim}{V}_{Nk}) > d^2Nk/(N-1)\chi^2_{p-1,\alpha}$. When $k = n_d^*$, sampling is terminated and $I_{Nn_d^*}$, defined by (4.9) for $n = n_d^*$, is taken as a (simultaneous) confidence region for $\underset{\sim}{\beta}_N$. Note that if $\underset{\sim}{A}$ be any p.d. matrix, then by the Schwarz inequality

$$\begin{aligned} \sup\{|\underset{\sim}{\ell}'\underset{\sim}{x}|: \underset{\sim}{\ell}'\underset{\sim}{\ell} = 1\} &= \sup\{|\underset{\sim}{\ell}'\underset{\sim}{A}^{\frac{1}{2}}\underset{\sim}{A}^{-\frac{1}{2}}\underset{\sim}{x}|: \underset{\sim}{\ell}'\underset{\sim}{\ell} = 1\} \\ &\leq [\sup\{\underset{\sim}{\ell}'\underset{\sim}{A}\underset{\sim}{\ell}: \underset{\sim}{\ell}'\underset{\sim}{\ell} = 1\}(\underset{\sim}{x}'\underset{\sim}{A}^{-1}\underset{\sim}{x})]^{\frac{1}{2}} \\ &= [ch_1(A)\ (x'A^{-1}x)]^{\frac{1}{2}} . \qquad (4.11) \end{aligned}$$

Hence, choosing $\underset{\sim}{A} = \underset{\sim}{V}_{Nn_d^*}$ and $\underset{\sim}{x} = (\underset{\sim}{b}^*_{Nn_d^*} - \underset{\sim}{\beta}_N)$, we obtain from (4.9), (4.10) and (4.11) that the maximum diameter of $I^*_{Nn_d^*}$ is

$$2\{ch_1(\underset{\sim}{V}_{Nn_d^*})[(N-n_d^*)\chi^2_{p-1,\alpha}/Nn_d^*)\}^{\frac{1}{2}} \leq 2d , \quad (4.12)$$

so that the width of the confidence interval for any $\underset{\sim}{\ell}'\underset{\sim}{\beta}_N$ is bounded by $(\underset{\sim}{\ell}'\underset{\sim}{\ell})^{\frac{1}{2}}\, 2d$, i.e., $I_{Nn_d^*}$ is a bounded-diameter confidence region of $\underset{\sim}{\beta}_N$. We intend to show that a d is chosen small,

$$P\{\underset{\sim}{\beta}_N \in I_{Nn_d^*}|\underset{\sim}{\beta}_N\} \to 1-\alpha , \quad (4.13)$$

insuring that the confidence coefficient of $I_{Nn_d^*}$ approaches $1-\alpha$ when d is chosen small. Towards this, we define

$$n_d^0 = \min\{k: k \geq n^0 \text{ and } ch_1(\underset{\sim}{\Gamma}_N) \leq d^2kN/(n-k)\chi^2_{p-1,\alpha}\} . \quad (4.14)$$

Then, by (3.15), (4.10) and (4.14), we have

$$n_d^*/n_d^0 \to 1 \text{ s.c., as } d \downarrow 0 , \quad (4.15)$$

and as a result, using the continuity (or, rather, the tightness)

property of $\underset{\sim}{Z}_N$ [implied by the convergence in law of $\underset{\sim}{Z}_N$ to $\underset{\sim}{Z}^0$], we obtain that $\underset{\sim}{Z}_N(N^{-1}n_d^*) - \underset{\sim}{Z}_N(N^{-1}n_d^0) \overset{p}{\to} 0$, while by Theorem 3.4, $\underset{\sim}{Z}_N(N^{-1}n_d^0)$ has asymptotically a multivariate normal distribution. Combining this with (3.15), we conclude that

$$T_{Nn_d^*} - T_{Nn_d^0} \to 0 \quad \text{s.c., as } d \downarrow 0 \ , \tag{4.16}$$

and using (4.16) along with the fact that $T_{Nn_d^0}$ has asymptotically the chi-square distribution with $p-1$ DF, we conclude that

$$\begin{aligned} P\{\underset{\sim}{\beta}_N \in I_{Nn_d^*} | \underset{\sim}{\beta}_N\} &= P\{T_{Nn_d^*} \geq \chi^2_{p-1,\alpha} | \underset{\sim}{\beta}_N\} \\ &\to P\{T_{Nn_d^0} \leq \chi^2_{p-1,\alpha} | \underset{\sim}{\beta}_N\} \to 1-\alpha \quad \text{as } d \downarrow 0 \ . \end{aligned} \tag{4.17}$$

Thus, (4.13) holds. The theory developed here is an extension of the Chow-Robbins (1965) theory of fixed-width confidence intervals to finite population sampling and to a more general class of statistics.

In medical trials, often, repeated significance tests (RST) are made on an increasing sequence of sample sizes with a view to stopping earlier if a significant result is obtained at that time (prior to reaching the target sample size). Here, we shall develop such RST procedures for sampling from a finite population. The theory rests on the invariance principles studied in Section 3. Let

$$T_{Nk} = (k^2/N)(\underset{\sim}{b}^*_{Nk} - \underset{\sim}{\beta}_N)' V^{-1}_{Nk} (\underset{\sim}{b}^*_{Nk} - \underset{\sim}{\beta}_N) \quad \text{for } N^* \leq k \leq N \ , \tag{4.18}$$

where $N^* \to \infty$ but $N^{-1}N^* \to 0$ as $N \to \infty$, while we let $T_k = 0$, $\forall\ k < N^*$. Let then

$$K_{Nn} = \max_{k \leq n} T_{Nk} \quad \text{and} \quad M_{Nn} = N^{-1} \textstyle\sum_{k<n} T_{Nk} \ , \tag{4.19}$$

where $n/N \to \nu \in (0,1]$. Then, we have

$$K_{Nn} \overset{\mathcal{D}}{\to} \sup_{0 \leq t \leq \nu} \left\{ \sum_{j=1}^{p-1} W^2_{j0}(t) \right\} = K_\nu^{(p-1)} \ , \quad \text{say} \ , \tag{4.20}$$

$$M_{Nn} \overset{\mathcal{D}}{\to} \sum_{j=1}^{p-1} \int_0^\nu W^2_{j0}(t)dt = W_\nu^{(p-1)} \ , \quad \text{say} \ ,$$

where $\{W_{j0}(t),\ t \in I\}$ are independent copies of a standard

Brownian bridge on I. Let $K_{\nu,\alpha}^{(p-1)}$ and $M_{\nu,\alpha}^{(p-1)}$ be the upper $100\alpha\%$ point of the distributions of $K_{\nu}^{(p-1)}$ and $M_{\nu}^{(p-1)}$, respectively. [For $p = 2$, $K_{\nu}^{(1)}$ and $M_{\nu}^{(1)}$ are functionals of a single Brownian bridge and their distributions are known; see Koziol and Byar (1975) and Pettitt and Stephens (1976). For $p \geq 3$, analytical solutions appear to be intractable; however, the prospect of simulation is quite bright. We may refer to Majumdar (1976) for some related work.]

Suppose now that we desire to test H_0: $\underset{\sim}{\beta}_N = \underset{\sim}{\beta}_0$ (specified) against $\underset{\sim}{\beta}_N \neq \underset{\sim}{\beta}_0$. Let T_{Nk}^0 be defined by (4.18) when $\underset{\sim}{\beta}_N$ is replaced by $\underset{\sim}{\beta}_0$ and M_{nn}^0 (and K_{Nn}^0) be defined by (4.19) when T_{Nk} is replaced by T_{Nk}^0, $k \geq 1$. Then, we have the following RST procedure:

Continue sampling as long as $k \leq n$ and T_{Nk}^0 (or M_{Nk}^0) is $< K_{\nu,\alpha}^{(p-1)}$ (or $M_{\nu,\alpha}^{(p-1)}$). If, for the first time, for $k = D$ ($\leq n$), $T_{D0}^2 \geq K_{\nu,\alpha}^{(p-1)}$ (or $M_{ND}^0 \leq M_{\nu,\alpha}^{(p-1)}$), stop sampling when $\underset{\sim}{X}_{ND}$ is observed, along with the rejection of H_0. If, no such $D(\leq n)$ exists, stop sampling at the preplanned n-th stage (i.e., when $\underset{\sim}{X}_{Nn}$ is observed), along with the acceptance of H_0.

By (4.9) through (4.13), we conclude that the asymptotic level of significance of this test is equal to α. Also, $E(D) \leq n$, indicating a saving in the average amount of sampling over the fixed-sample size procedure. In fact, we may even test for a more general hypothesis:

$$H_0: \underset{\sim}{C}\underset{\sim}{\beta}_N = \underset{\sim}{\beta}_0 \quad \text{vs.} \quad H: \underset{\sim}{C}\underset{\sim}{\beta}_N \neq \underset{\sim}{\beta}_0 , \tag{4.22}$$

where $\underset{\sim}{C}$ is a $q \times (p-1)$ matrix of rank $q (1 \leq q \leq p-1)$ and $\underset{\sim}{\beta}_0$ is a specified q-vector. For this case, in (4.18), we need to take $T_{Nk} = N^{-1}k^2(\underset{\sim}{b}_{Nk}^{*}\underset{\sim}{C}' - \underset{\sim}{\beta}_0')(\underset{\sim}{C}\underset{\sim}{V}_{Nk}\underset{\sim}{C}')^{-1}(\underset{\sim}{C}\underset{\sim}{b}_{Nk}^{*} - \underset{\sim}{\beta}_0)$, $k \geq N^*$ (and equal to 0 for $k < N^*$), and in (4.20)-(4.21), we need to change $p-1$ to q. Rest of the sequential procedure remains the same.

4.2. Estimation and Testing of Ratio of Means.

In the same set up of Section 4.1, we consider the parameters

$$\rho_{N(ij)} = \bar{a}_N^{(i)}/\bar{a}_N^{(j)} \ , \quad \text{for} \quad 1 \le i < j \le p \ , \tag{4.23}$$

and our interest centers around one or more of these ratios. The usual estimator of $\rho_{N(ij)}$ is

$$\hat{\rho}_{N(ij)} = \bar{X}_{Nn}^{(i)}/\bar{X}_{Nn}^{(j)} \ , \quad \bar{X}_{Nn} = n^{-1}\sum_{i=1}^{n} X_{Ni}$$
$$= (\bar{X}_{Nn}^{(1)},\ldots,\bar{X}_{Nn}^{(p)}) \ , \tag{4.24}$$

for $1 \le i < j \le p$. Though $\bar{X}_{Nn}$ is a U-statistics (vector) and is unbiased for $\bar{\underset{\sim}{a}}_N$, $\hat{\rho}_{N(ij)}$ is not generally an unbiased estimator of $\rho_{N(ij)}$. Hence, jackknifing may be employed to reduce the bias. Here, we note that

$$(\partial/\partial b_i)(b_1/b_2) = (-1)^{i-1}(b_1/b_2)^{i-1}b_2^{-1} \quad \text{for} \quad i = 1,2 \ . \tag{4.25}$$

Hence, we may proceed as in (4.9) through (4.22) and consider point as well as confidence interval estimators of $\rho_{N(ij)}$, $1 \le i < j \le p$ and also (sequential or nonsequential) tests for any subset of these parameters. In passing, we may remark that the Grizzle-Starmer-Koch (1969) linear models with categorical data extend to the situation when proportions are ratios and samples are drawn without replacement from finite populations, and where jackknifing is employed for bias reduction. This is actually a special case of (4.23) when each of the p arguments of $\underset{\sim}{a}_{Ni}$ is either 0 or 1, so that $\bar{a}_N^{(i)}$ is the proportion of 1's in the N responses on the i-th characteristic, for $i = 1,\ldots,p$.

4.3. Optimal Allocation in Stratified Random Sampling

Suppose that the population of N units is subdivided into $r(\ge 2)$ sub-populations of sizes $N_1,\ldots,N_r$ (so that $N = N_1 + \ldots + N_r$). Let $\bar{\underset{\sim}{a}}_{Nh}$ and $\underset{\sim}{\Lambda}_{Nh}$ be defined as in (4.1) for the h-th sub-population, $h = 1,\ldots,r$. Suppose that a sample of size n is to be drawn and let $n_1,\ldots,n_r$ denote the sub-sample sizes for the r strata. Optimal allocation of $n_1,\ldots,n_r$ [viz.,

Chapter 5 of Cochran (1963)] depends on $\underset{\sim}{\Lambda}_{N1},\ldots,\underset{\sim}{\Lambda}_{Nr}$, which are all unknown. We may start with an initial sample of size n_0 $(= r\bar{n}_0)$ with $\bar{n}_0$ observations from each stratum, estimate the $\underset{\sim}{\Lambda}_{Nh}$, $1 \le h \le r$ and using these estimates get an estimated optimal allocation for n; this usual practice entails some loss of efficiency. As in Williams and Sen (1973), we may consider a multi-stage (or sequential procedure) where we keep on updating the estimators of $\underset{\sim}{\Lambda}_{Nh}$, $1 \le h \le r$, so that the procedure will be asymptotically optimal. In this context, jackknifing can also be used—the theorems studied in Section 3 insure that for jackknifing, the sequential procedure leads to an asymptotically optimal allocation.

5. PROOF OF THEOREM 3.2

We consider here the proof of Theorem 3.2. Note that by (2.11),

$$U_{Nn} - \mu_N = mW_{Nn}^{(1)} + W_{Nn}^{*} \ ; \ W_{Nn}^{*} = \sum_{h=2}^{m} \binom{m}{h} W_{Nn}^{(h)} \ , \qquad (5.1)$$

for every $N \ge n \ge m$. Hence, for every $n > m$,

$$U_{Nn-1} - U_{Nn} = m(W_{Nn-1}^{(1)} - W_{Nn}^{(1)}) + (W_{Nn-1}^{*} - W_{Nn}^{*}) \ . \qquad (5.2)$$

By virtue of (2.5), (2.9), (2.10) and the results of Nandi and Sen (1963), it follows by some routine steps that

$$E\{[n(n-1)(W_{Nn-1}^{*} - W_{Nn}^{*})^2]^2\} \le cn^{-2} \ , \ \forall \ m < n \le N \ , \qquad (5.3)$$

where c does not depend on N. Hence, for every $N^{*}(\le N)$: $N^{*} \to \infty$,

$$P\left\{\max_{N^{*}\le n\le N} E[n(n-1)(W_{Nn-1}^{*} - W_{Nn}^{*})^2 | C_{Nn}] > \varepsilon\right\}$$

$$\le \sum_{n=N^{*}}^{n} P\{E[n(n-1)(W_{Nn-1}^{*} - W_{Nn}^{*})^2 | C_{Nn}] > \varepsilon\}$$

$$\le \sum_{n=N^{*}}^{N} \varepsilon^{-2} E[n(n-1)(W_{Nn-1}^{*} - W_{Nn}^{*})^2]^2$$

$$\le 2\varepsilon^{-2} \sum_{n=N^{*}}^{N} n^{-2} \le c\varepsilon^{-2}(N^{*}-1)^{-1} \to 0 \ , \ \forall \ \varepsilon > 0 \ , \qquad (5.4)$$

Hence, to prove the theorem, it suffices to replace $n(n-1)$ $(U_{Nn-1}-U_{Nn})^2$ by $m^2 n(n-1)(W_{Nn-1}^{(1)}-W_{Nn}^{(1)})^2$. Toward this, not that

$$n(n-1)E\{(W_{Nn-1}^{(1)} - W_{Nn}^{(1)})^2 | C_{Nn}\}$$

$$= n(n-1)^{-1}E\{[W_{Nn}^{(1)} - \{f_1(X_{Nn}) - \mu_N\}]^2 | C_{Nn}\}$$

$$= n(n-1)^{-1}[n^{-1}\sum_{i=1}^{n}\{f_1(X_{Ni}) - \mu_N\}^2 - \{W_{Nn}^{(1)}\}^2] . \qquad (5.5)$$

Note that $\{W_{Nn}^{(1)}, C_{Nn}, n \geq m\}$ is a reverse martingale (being a U-statistic sequence) and $E\{W_{Nn}^{(1)}\}^2 = \bar{\zeta}_{1,N}((N-n)/Nn)$. Hence, by Theorem 1 of Sen (1970), $\forall\ \varepsilon > 0$

$$P\left\{\max_{N^* \leq n \leq N} [W_{Nn}^{(1)}]^2 > \varepsilon\right\} \leq \varepsilon^{-1}\bar{\zeta}_{1,N}\{(N-N^*)/NN^*\}$$

$$\to 0 \quad \text{as} \quad N^* \to \infty . \qquad (5.6)$$

On the other hand, $\tilde{U}_{Nn} = n^{-1}\sum_{i=1}^{n}\{f_1(X_{Ni}) - \mu_N\}^2$ $(n = 1)$ is also a U-statistic with $E\tilde{U}_{Nn} = \bar{\zeta}_{1,N}$ and by (2.5) and by (3.14) of Nandi and Sen (1963), $V\{\tilde{U}_{Nn}\} = 0(n^{-1}(N-n)/N)$, $\forall\ n \leq N$. Hence, by Theorem 1 of Sen (1970), we have for every $\varepsilon > 0$,

$$P\left\{\max_{N^* \leq n \leq N} |\tilde{U}_{Nn} - \bar{\zeta}_{1,N}| > \varepsilon\right\} \leq \varepsilon^{-2}V(\tilde{U}_{NN^*})$$

$$\to 0 \quad \text{as} \quad N^* \to \infty . \qquad (5.7)$$

From (5.5), (5.6), and (5.7), it follows that $|n(n-1)E\{(W_{Nn-1}^{(1)} - W_{Nn}^{(1)})^2 | C_{Nn}\} - \bar{\zeta}_{1,N}| \to 0$ s.c., and the proof of the theorem is complete.

Remark. For infinite populations, a similar result has been proved by Bhattacharyya and Sen (1977). In view of the relatively stringent assumption (2.5), for finite population sampling, the present proof is considerably simpler in nature.

ACKNOWLEDGEMENTS

The work was partially supported by the Air Force Office of Scientific Research, A.F.S.C., U.S.A.F., Grant No. AFOSR-74-2736 and partially by the National Heart, Lung and Blood Institute, Contract NIH-NHLBI-71-2243 from the NIH.

BIBLIOGRAPHY

Arvesen, J.N. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* 40, 2076-2100.

Bhattacharyya, B.B. and Sen, P.K. (1977). Weak convergence of the Rao-Blackwell estimator of a distribution function. *Ann. Prob.* 5, 500-10.

Chow, Y.S. (1960). A martingale inequality and the law of large numbers. *Proc. Amer. Math Soc.* 11, 107-11.

Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed with sequential confidence intervals for the mean. *Ann. Math. Statist.* 36, 457-462.

Cochran, W.G. (1963). *Sampling Techniques* (2nd. ed.). New York: John Wiley & Sons, Inc.

Grizzle, J.E., Starmer, C.F., and Koch, G.G. (1969). Analysis of categorical data by linear models. *Biometrics* 25, 489-504.

Koziol, J.A. and Byar, D.P. (1975). Percentage points of the asymptotic distribution of the one and two sample K-S tests for truncated or censored data. *Technometrics* 17, 507-10.

MacRae, E.C. (1974). Matrix derivatives with an application to an adaptive linear decision problem. *Ann. Statist.* 2, 337-46.

Majumdar, H. (1976). Generalized rank tests for progressive censoring procedures. Unpublished Ph.D. dissertation. Inst. Statist. Mimeo Report #1081, Univ. North Carolina, Chapel Hill.

Nandi, H.K. and Sen, P.K. (1963). On the properties of U-statistics when the observations are not independent. Part two: Unbiased estimation of the parameters of a finite population. *Calcutta Statist. Assoc. Bull.* 12, 124-48.

Pettitt, A.N. and Stephens, M.A. (1976). Modified Cramér-von Mises statistics for censored data. *Biometrika* 63, 291-98.

Rosén, B. (1964). Limit theorems for sampling from a finite population. *Arch. Mat.* 5, 383-424.

Rosén, B. (1967a). On the central limit theorem for the sums of dependent random variables. *Zeit. Wahrscheinlichkeitsth* 7, 47-82.

Rosén, B. (1967b). On the central limit theorem for a class of sampling procedures. *Zeit. Wahrscheinlichkeitsth* 7, 103-15.

Sen, P.K. (1970). The Hájek-Rényi inequality for sampling from a finite population. *Sankhyā A* 32, 181-88.

Sen, P.K. (1972). Finite population sampling and weak convergence to a Brownian bridge. *Sankhyā A* 34, 85-90.

Sen, P.K. (1977). Some invariance principles relating to jack-knifing and their role in sequential analysis. *Ann. Statist.* 5, 316-29.

Williams, G.W. and Sen P.K. (1973). Asymptotically optimal sequential estimation of regular functionals of several distributions based on generalized U-statistics. *J. Multivariate Analysis* 3, 469-82.

Key Words and Phrases: Brownian bridge, optimal allocation, random sampling without replacement, repeated significance tests, sequential confidence regions, strong convergence, Tukey estimator.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 78-1303	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) INVARIANCE PRINCIPLES FOR JACKKNIFING U-STATISTICS FOR FINITE POPULATION SAMPLING AND SOME APPLICATIONS		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Hiranmay Majumdar and Pranab Kumar Sen		8. CONTRACT OR GRANT NUMBER(s) AFOSR-74-2736
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of North Carolina Department of Biostatistics Chapel Hill, North Carolina 27514		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5 A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE 1978
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

12 22p.

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Brownian bridge, optimal allocation, random sampling without replacement, repeated significance tests, sequential confidence regions, strong convergence, Tukey estimator.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

For simple random sampling (without replacement) from a finite population, suitable stochastic processes are constructed from the entire sequence of jackknife estimators based on functions of U-statistics and these are approximated in distribution by some Brownian bridge processes. Strong convergence of the Tukey estimator of the variance of jackknife U-statistics has also been established. Some applications of these results in sequential analysis relating to finite population sampling are also considered.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408 985